Chunk hierarchies and retrieval structures: Comments on Saariluoma and Laine

Gobet, F. (2001). Chunks hierarchies and retrieval structures. Comments on Saariluoma and Laine. *Scandinavian Journal of Psychology, 42*, 149-155.

Fernand Gobet
ESRC Centre for Research in
Development, Instruction and Training
Department of Psychology
University of Nottingham
Nottingham NG7 2RD
England

Phone: (0115) 951 5402
Fax: (0115) 951 5324
frg@psyc.nott.ac.uk

Running head: Chunk hierarchies and retrieval structures

Abstract

The empirical results of Saariluoma and Laine (in press) are discussed and their computer simulations are compared with CHREST, a computational model of perception, memory and learning in chess. Mathematical functions such as power functions and logarithmic functions account for Saariluoma and Laine's (in press) correlation heuristic and for CHREST very well. However, these functions fit human data well only with game positions, not with random positions. As CHREST, which learns using spatial proximity, accounts for the human data as well as Saariluoma and Laine's (in press) correlation heuristic, their conclusion that frequency-based heuristics match the data better than proximity-based heuristics is questioned. The idea of flat chunk organisation and its relation to retrieval structures is discussed. In the conclusion, emphasis is given to the need for detailed empirical data, including information about chunk structure and types of errors, for discriminating between various learning algorithms.

Acknowledgements

I thank Pekka Niemi for inviting me to write this comment, Pertti Saariluoma and Tei Lane for sending me the human data and the positions used in their simulations, as well as Peter Lane and Julian Pine for comments on this manuscript.

Chunk hierarchies and retrieval structures: Comments on Saariluoma and Laine (in press)

Saariluoma and Laine (in press) present interesting data on early learning in skill acquisition, and explore computational mechanisms allowing such learning. They also make the theoretical proposal that chunks have a flat organisation that can be used as a retrieval structure, which contrasts with the hierarchical organisation typically used for that purpose in cognitive psychology.

In this paper, I first comment on Saariluoma and Laine's experimental results and on their modelling experiments. I then discuss their theoretical proposal of a flat organisation of chunks, and compare it with the organisation used by CHREST,[1] a computational theory of chess expertise (De Groot & Gobet, 1996; Gobet, 1993; Gobet & Simon, 1996b; in press). Simulations with CHREST of the data presented by Saariluoma and Laine will be used to illustrate some of the differences between the two approaches.

The Experimental Data

While a great deal is known about learning in simple tasks (especially perceptual learning, cf. Shiffrin, 1996) and about expertise in general, relatively little is known about learning at the early stages of expertise. With chess, there is little more about this topic than the experiments carried out by Ericsson and Harris (1990) and Fisk and Lloyd (1988). Therefore, the experiment described by Saariluoma and Laine, where two subjects were trained to improve their memory for briefly-presented chess positions, represents a welcome addition to the literature.

---

[1]CHREST stands for Chunk Hierarchy and REtrieval STructures.

Two main results may be singled out from this experiment. First, with game positions, both subjects improved rapidly at the beginning, and then increasingly more slowly, a curve that is captured by logarithmic or power functions.[2] Second, both subjects showed a slight increase in the recall of random positions. Both results are predicted by theories explaining expertise by the regular acquisition of chunks (for the power law of learning, see Newell & Rosenbloom, 1981; for the use of logarithmic functions describing chunking-based learning and for the recall of random positions, see Gobet & Simon, 1996a, 1996b).

Although these data are interesting, it is a pity that Saariluoma and Laine do not give more details about strategies, errors, or type of chunks replaced by the subjects, as was for example done in Chase and Ericsson's (1982) seminal research on digit-span memory. Indeed, Ericsson and Harris' (1990) sister study on training a novice in recalling chess positions would at least suggest that Saariluoma and Laine's subjects did use some kind of explicit strategy during the learning phase. Information about how well the positions in the training set were learnt would have been of interest as well. Such details would have given more diagnostic power to the data, in particular with respect to modelling. Another slight problem with the data is that both subjects had some knowledge of chess at the beginning of the experiment, which probably inflates the estimate of their recall performance. However, despite these minor problems, the data are useful for exploring various learning mechanisms, which is how they are used by Saariluoma and Laine, or for testing the generality of an existing cognitive architecture, which is what I will do with CHREST later in this paper.

---

[2]Psychologists prefer to use power functions ($y = a * x^b$, where b is typically < 1 and > -1) because they work well both in the case where y increases with x (e.g., percentage correct as a function of practice) and in cases where y decreases with x (e.g., reaction time as a function of practice). Logarithmic functions ($y = a \pm b * \log [x]$) work well in the first case, but are not plausible in the second case, because they do not reach asymptote (in our second example, a log function would predict negative reaction time!).

As noted above, both subjects improved their recall performance with random positions. Indeed, it is of special interest that Saariluoma and Laine used random positions in their experiment, as recent work has shown that these positions do not simply constitute a control task, as was thought for a long time, but are highly discriminative for theories of expertise. It is well known that Chase and Simon (1973a; 1973b) found that skill effects disappear with briefly-presented random positions, a result that led to an intriguing situation. On the one hand, this finding has often been heralded as one of the most robust phenomena in expertise research (e.g., Ericsson & Kintsch, 1995; Gobet, 1993, Holding, 1985; Saariluoma, 1995). On the other hand, some issues did not fit the puzzle quite so well. For example, Chase and Simon found that "even in the randomised boards, players are noticing the same kinds of structures as those they perceive in the coherent positions, even though these structures occur rarely in randomised boards" (1973b, p. 232), which seems to suggest some kind of skill effect. The presence of adventitious chunks in random positions has been noted by several authors (Chase & Simon, 1973a; 1973b; Saariluoma, 1984, 1989; Simon & Barenfeld, 1973).

It is only recently that the riddle was solved. Gobet and Simon (1996a), combining the result of all published studies they could find on the recall of random chess positions, showed that there is a reliable skill effect, albeit a small one. This effect, however, was not significant in most individual experiments, due to their lack of statistical power. Gobet and Simon (1996a, 1996b) also showed that the chunking theory (Chase & Simon, 1973a, 1973b; Simon & Gilmartin, 1973) and related models predict such a (slight) skill superiority in random positions, because the presence of a large database of chunks makes it more likely to find at least a few chunks in a random position. More recently, Gobet and Simon (in press) showed that CHREST makes accurate quantitative predictions on the recall of random positions, with presentation times spanning from one second to one minute. The skill effect with random positions is theoretically important, because it is difficult to explain with theories based on high-level concepts, such as the theories proposed by Cooke et al.

(1993) and by Holding (1985). Here, I certainly agree with Saariluoma and Laine that chunks play a key role in chess expertise, and probably in most other kinds of expertise.

The Computer Simulations

In their simulations, Saariluoma and Laine contrast two chunking heuristics. First, a heuristic where chunks are built around a focal piece, using adjacent pieces (random neighbourhood heuristic). Second, a heuristic based on the frequency of co-occurrence and similarity of pieces, which is not constrained by spatial proximity (correlation heuristic). The correlation heuristic fits the data reasonably well. However, a surprising feature of the simulations is that, with both game and random positions, the random neighbourhood heuristic actually gets worse as it learns additional chunks. It is not really clear from the description of the heuristics given by Saariluoma and Laine why this should be so. It is also unclear why the two versions differ at the beginning of the experiment, when no learning has taken place.

As noted by Saariluoma and Laine, the correlation heuristic learns in a way reminiscent of neural nets. A consequence is that it makes predictions about the kind of chunks learnt that differ from the predictions of the random neighbourhood heuristic as well as of the chunking theory and CHREST, which emphasise spatial proximity in learning. What kinds of chunks chess players really learn could be tested in experiments where, for example, the type of chunks acquired by either method were flashed for a few seconds on a computer screen, and recall performance was assessed.

Saariluoma and Laine's goal was not really to run cognitive simulations, but to compare two learning methods. Even so, in order to understand chunking in general, it is worth mentioning some features of their simulations which do not match the empirical data. Contrary to the human data (Chi, 1978) there is no overlap between chunks. Nor does the program make any errors. Finally, the assumption that the program starts with a chunk for each combination of piece and square (a total of 768 chunks) leads to a recall which is too high with random positions at the beginning of learning. While it is probably true that novices can distinguish different kinds of pieces on different squares, as argued by Saariluoma and Laine, it is unlikely that they can memorise them—which is what the program is doing.

Newell and Rosenbloom (1981) showed that chunking mechanisms acquiring new chunks at a constant rate yield performance improvements that follow a power function, which Saariluoma and Laine prefer to refer to as a logarithmic function (as we shall see below, empirical data are often not powerful enough to tease the two functions apart). Simulations with the correlation heuristic confirm Newell and Rosenbloom's analysis, as do simulations with CHREST (see below; see also Gobet & Simon, 1996b). By contrast, the simulations with the random neighbourhood heuristic do not follow this pattern. This is quite surprising, as Newell and Rosenbloom's analysis was done at a rather high level of abstraction, and would seem to cover a large class of chunking algorithms. It is unclear from Saariluoma and Laine's description why this "anomaly" occurs. Another departure from Newell and Rosenbloom's analysis is that the latter suggest that a single mechanism is enough for yielding negatively accelerating learning, while Saariluoma and Laine claim that several mechanisms are necessary (cf. Saariluoma & Laine, p. 5).

Chunk Hierarchies and Retrieval Structures

The main theoretical contribution of Saariluoma and Laine's paper is their discussion of the organisation of chunks and its relation to retrieval structures. Models in the EPAM family organise chunks as a hierarchy, which develops dynamically as a function of learning (see Feigenbaum & Simon, 1984, or De Groot & Gobet, 1996, for details on the learning algorithms). By contrast, Saariluoma and Laine propose a flat, modular organisation, which is similar to that used by most production systems (Newell & Simon, 1972). Both representations are plausible (as many others), and it is unclear whether current empirical data can discriminate between them.

The next step taken by Saariluoma and Laine is intriguing, however. They first define retrieval structure as meaning that "subjects have some pieces of knowledge in their immediate working memory and the rest of task-relevant information is stored into long-term memory [...]" (p. 8). They then propose that "[the] retrieval structure is formed by a set of parallel and non-integrated patterns" and that "the contents of the patterns themselves cause the integration but no direct links combining patterns are required" (Saariluoma and Laine, in press, p. 8). Before commenting on this proposal, it is necessary to briefly review how the concept of retrieval structure has been used in recent research.

Retrieval structures have enjoyed great popularity in recent years as a means of accounting for (expert) memory in various domains. However, one difficulty with this concept is that different authors use it with different meanings. Chase and Ericsson (1982), who originated the term, give the following definition: "A retrieval structure is a long-term memory structure for indexing material in long-term memory. It can be used to store and order information, but is more versatile because it can allow direct retrieval of any identifiable location. A good example of a retrieval structure is the mnemonic system known as the Method of Loci [...]" (p. 17). Note the importance given to storing information, not only to retrieving it. Chase and Ericsson also emphasise the hierarchical structure of retrieval structures and the fact that it takes a massive amount of practice to learn them.

Ericsson and Kintsch (1995) use a similar definition: Retrieval structures are "a set of retrieval cues [that] are organized in a stable structure" (Ericsson & Kintsch, 1995, p. 216; see also their Figure 1). They also stress that, in addition to retrieval structures, "knowledge-based associations relating units of encoded information to each other along with patterns and schemas [...]" (p. 221) are necessary for expert memory.

Kintsch (1998, p. 74) proposes a model where "knowledge is represented as a network of propositions. Such a network is called a *knowledge net*. The nodes of the net are propositions, schemas, frames, scripts, production rules [...]." Retrieval structures are defined within this framework: "[...] most nodes in a knowledge network are connected with powerful, stable links—retrieval structures—to other nodes in the net that can be brought into working memory" (p. 74). Kintsch's definition is more encompassing than Ericsson and Kintsch's (1995), since it includes patterns and schemas, which are clearly treated separately by Ericsson and Kintsch from cue-based retrieval (retrieval structures), although the two types of encoding are supposed to interact (compare Figure 4 of Ericsson & Kintsch, 1995, with Figure 7.2 of Kintsch, 1998).

Like Chase and Ericsson (1982), Richman, Staszewski and Simon (1995) as well as Gobet and Simon (1996c; in press), who work in the EPAM tradition, emphasise that retrieval structures can store information swiftly and they have a hierarchical organisation. For them, the key aspect of a retrieval structure is that it contains slots (variables) that allow values to be encoded rapidly. They distinguish between structures acquired *explicitly* to meet the demands of the task (retrieval structure in the strict sense), for example by the subjects trained in the digit-span task, and structures acquired *implicitly* in the acquisition of expert knowledge, for example by chess players. They call the latter structures "templates". Both types of retrieval structure have been implemented as computer programs simulating expert memory in the digit-span task and in chess respectively. Note that these authors do not consider an index to long-term memory (LTM) such as the EPAM discrimination net as a retrieval structure, because it lacks the property of allowing rapid encoding. In this respect, their use of the term is consistent with that of Chase and Ericsson (1982, p. 16-17).[3]

---

[3] Perhaps, some confusion could have been avoided about the concept of retrieval structure by using two terms, such as "storage structures" for structures allowing rapid storage and facilitating retrieval, and "retrieval structures" for structures only facilitating retrieval but not allowing rapid storage. Another source of confusion is that this concept is often used in a way that includes both the structure itself and the mechanisms associated with it.

As can be seen, Saariluoma and Laine's definition is clearly new: it does not include the ideas of rapid storage, of hierarchical organisation, and of linked nodes, present in all the sources just mentioned. In a sense, it is even less restrictive than Kintsch's (1998) definition, which still provided structure through the links connecting nodes in semantic memory. As noted by Saariluoma and Laine, it is probably too early to evaluate their framework, as more detail about the psychological mechanisms allowing information to be learnt and retrieved are lacking. Perhaps the lack of a hierarchical structure is debatable, since, as we have just seen, it has often been emphasised in the literature on expertise. In addition, the results of Freyhof, Gruber and Ziegler (1992), who asked chessplayers to partition positions at various levels of granularity, would seem to support a hierarchical organisation.

Obviously, verbal arguments will not settle the question. Saariluoma and Laine present computer simulations, and it is incumbent on theorists defending a different point of view to offer such simulations as well. Fortunately, the template theory (Gobet & Simon, 1996c; in press) is implemented as a program, called CHREST, which incorporates both the idea of a hierarchical organisation of chunks and the concept of retrieval structure. It is therefore possible to compare it in detail to the theoretical ideas advanced by Saariluoma and Laine.

CHREST: A Computational Theory of Chess Expertise

CHREST (De Groot & Gobet, 1996; Gobet, 1993; Gobet & Simon, 1996b; in press) is a computational theory of chess expert perception, memory and learning. It is inspired by Simon and Barenfeld's (1969) and Simon and Gilmartin's (1973) programs of perception and memory in chess. CHREST consists of a short-term memory (STM), limited to four items, of a discrimination net, and of mechanisms directing eye movements and managing memory. It acquires chunks (symbols denoting patterns of pieces on the board) by growing a discrimination net when scanning positions from a database of chess games. The discrimination net provides a structure in which chunks are organised hierarchically.

Some of the large chunks evolve into more complex data structures, called templates, which contain a core (similar to chunks) and variable slots, where the information can be rapidly stored. Templates are essentially schemata; the originality of CHREST is in providing mechanisms on how schemata can be learnt, how they relate to perceptual information, and in providing estimates of the time it takes to encode information into schemata. A version of CHREST plays (weak) chess by pure pattern recognition, where recognised patterns elicit potential moves (Gobet & Jansen, 1994).

Eye movements play an important role in CHREST, as they determine the focus of attention, and attention in turn determines what will be learnt. Six mechanisms potentially direct eye movements (De Groot & Gobet, p. 233-236): (a) LTM information; (b) perceptual salience; (c) lines of force (attack and defence); (d) random square in peripheral vision; (e) random piece in peripheral vision; and (f) heuristics aimed at gaining information from a part of the display that has not been fixated yet. Given that some of these mechanisms are not specific to chess, Saariluoma and Laine's (in press, p. 16) statement that "the major simulation models . . . have an important presupposition: they use only chess specific relations, and these are their only heuristics (Simon and Gilmartin 1973, de Groot and Gobet 1996)" is simply incorrect.

The program accounts for a large amount of data, including the role of presentation time in the recall of game and random positions (Gobet & Simon, in press), the effect of mirror-image modification of the board (Gobet & Simon, 1996b), eye movements during the first 5 seconds of the presentation of a position (De Groot & Gobet, 1996), and the recall of multiple boards (Gobet, 1993). Although most of the simulations have been done with the aim of explaining skilled behaviour, it seems worthwhile to see how well CHREST accounts for the data collected by Saariluoma and Laine on early learning. In fact, it turns out that CHREST does this rather well.

Methods

These simulations use the same version of CHREST as discussed by Gobet and Simon (in press); no parameter was altered in order to fit the data. The only difference was that eye-movement heuristics based on relations of defence and attack were disabled (see Saariluoma & Laine, p. 7).[4] To facilitate the comparison with Saariluoma and Laine's simulations, I followed their methodology : same database of positions; 500 positions in the learning set; tests after the study of 30, 60, 175, 220, 270, 350 and 500 positions; 10 game and 10 random test positions, presented for 5 seconds each; 20 independent runs. The only differences were as follows. First, advantage was taken of the fact that CHREST possesses detailed time parameters to carry out fine-grained simulations of the two human subjects, who studied each position for different amounts of time on average (roughly 3 minutes for NT and 5 minutes for MQ). One version of CHREST was allowed 3 minutes per position (CHREST-3), and another 5 minutes (CHREST-5). Second, the program started with zero chunks, but with some knowledge of the board and the pieces. With the recall tests at the beginning of the experiment, it was assumed that CHREST could memorise location, colour and type of piece in three different chunks, yielding a recall of one piece (cf. Gobet & Simon, 1996b). Third, only one estimate of STM capacity was used (three chunks). In previous simulations, Gobet (1998) varied STM capacity with the recall of random positions and found that this capacity gave a good fit to human performance. In addition, STM capacity has been estimated as about three items in domains other than chess (e.g., Zhang & Simon, 1985). Interestingly, this capacity is close to that which gives the best results in Saariluoma and Laine's simulations (four).

----

[4] Actually, simulations using defence and attack heuristics lead to similar results as those presented here.

Insert Figure 1 about here

Insert Table 1 about here

Results

Figure 1 shows that CHREST-3 and CHREST-5 account for the human data very well with game positions and reasonably well with random positions. Both human subjects do better at the beginning of the experiment than CHREST, which may be explained either by both having played some chess before the experiment or by a poor calibration of CHREST. Table 1 shows that the simulations are well fitted both by a power and a logarithmic function. The same applies to the human data with the recall of game positions. As is often the case with learning curves, it is very hard to distinguish between logarithmic functions and power law functions. Interestingly, the human data for the recall of random positions are not well captured by either function; Saariluoma and Laine propose a linear function to describe learning with the random positions, although a stepwise function would perhaps offer a better fit. The sudden jump present with both human subjects may indicate a change in strategy (Delaney, Reder, Staszewski & Ritter, 1998).

Figure 2 illustrates the distribution of chunks acquired by CHREST-5 after one run (a total of 13808 chunks). In this case, 708 templates were learnt. One can see that CHREST predicts that humans learn a considerable number of chunks and templates during this task. Recall performance remains modest (less than 45%), however, because the chunks and templates acquired are highly specialised and limited to the set of 500 positions learnt (in addition, these positions came from only 67 different games). By contrast, letting CHREST acquire about 10,000 chunks with a larger sample of positions studied each for a shorter time, as was done by Gobet and Simon (in press), allows recall performance of about 60% correct. It takes about 300,000 chunks to reach master performance.

Although CHREST does not necessarily use chess-specific information to direct eye movements, the chunks it learns reflect the type of interpiece relations found by Chase and Simon (1973a; see also Gobet & Simon, 1998). In particular, relations of colour are more frequent than relations of adjacency and sameness of kind, which in turn are more frequent than relations of defence. Relations of attack are the rarest of all. These properties emerge though the interaction of the learning environment (positions taken from actual games) and the chunking mechanism paying attention to spatial proximity. Finally, as Saariluoma and Laine do not give details of the type of chunks their subjects acquire and of the kind of errors they make, no such analysis will be provided here. Instead, the interested reader is referred to Gobet and Simon (in press).

Insert Figure 2 about here

Comparison with Saariluoma and Laine's Approach

In this section, CHREST will be evaluated with respect to the four points addressed by Saariluoma and Laine in their introduction and in their discussion. I will mostly limit the comparison to their correlation heuristic, as their random neighbourhood heuristic does not fit the human data very well.

First, the characteristics of the learning curve. Both CHREST and the correlation heuristic learn in a way that fits a negatively accelerating learning curve such as a power law or a logarithmic function. The human data supports such a learning curve with game positions, but not with random positions. Second, the nature of associative links. Saariluoma and Laine note that the heuristic based on frequency fits the data better than that based on proximity. CHREST, which learns in a way that emphasises spatial proximity, does not support this conclusion, as it fits the human data at least as well as the correlation heuristic. Obviously, frequency is also important with CHREST, since patterns that recur often in the training set are learnt faster than rare patterns. However, frequency is modulated by proximity in a way that is not achieved by the correlation heuristic. As seen above, CHREST learns chunks that reflect the properties found in human chunks. This is interesting in that relations of adjacency, colour, defence, and, more rarely, attack can be acquired with a learning algorithm that heeds only proximity (i.e., non-chess-specific information), assuming that the learning environment consists of coherent material. It is unclear whether the correlation heuristic produces such a result. At any rate, Saariluoma and Laine's claim about the superiority of the frequency heuristic is premature.

Third, the size of working memory. Previous simulations with CHREST have shown that a short-term memory capacity of three matches the human data well (Gobet, 1998; Gobet & Simon, in press), which is close to the number that gives the best fit in Saariluoma and Laine's simulations. CHREST is also consistent with their view that chess players construct retrieval structures to expand working memory—templates play that role. Finally, the question of hierarchy in retrieval structures. As noted above, Saariluoma and Laine's definition of retrieval structures would include the discrimination net used by CHREST and is therefore more inclusive than that of Simon, Richman, Staszewski and Gobet. The idea of a flat organisation of chunks is of interest, although it is unclear whether current data could discriminate between this kind of organisation and a hierarchical organisation.

Conclusion

Several points of disagreement, some important, have been uncovered in this paper. Some relate to methodological differences. For example, I prefer using richer data, including errors and information about chunks, as well as convergent evidence from various experiments, to evaluate computer models, while Saariluoma and Laine prefer comparing several learning methods on a single set of data (see Hyötyniemi & Saariluoma, 1998, or Laine, Hyötyniemi & Saariluoma, 1998, for a connectionist learning algorithm). Others are more substantial, for example my view that the hierarchical organisation of retrieval structures is critical, compared with Saariluoma and Laine's emphasis on a flat organisation, and my disagreement that their simulations are diagnostic of the superiority of frequency-based chunking over proximity-based chunking.

Although I have emphasised the differences between our approaches in this paper, it should be noted that they also share several similarities. In particular, there is clear agreement about the importance of chunking, the necessity of investigating retrieval structures, the strength of modelling for studying human cognition, and the need for theories of learning, including constructivist theories, to be formulated clearly and precisely, if possible as computer programs. It is unfortunately all too often the case that key concepts in psychology, such as "schema" and, as was shown here, "retrieval structure", evolve multiple meanings due to the flexibility offered by informal theorising. Even if they diverge on several points, our approaches to modelling at least offer the rigour necessary for elucidating these concepts.

Reference list

Chase, W. G., & Ericsson, K. A. (1982). Skill and working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation*, (Vol. 16,). New York: Academic Press.

Chase, W. G., & Simon, H. A. (1973a). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Chase, W. G., & Simon, H. A. (1973b). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, N.J.: Erlbaum.

Cooke, N. J., Atlas, R. S., Lane, D. M., & Berger, R. C. (1993). Role of high-level knowledge in memory for chess positions. *American Journal of Psychology, 106*, 321-351.

de Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess. Heuristics of the professional eye*. Assen: Van Gorcum.

Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (1998). The strategy specific nature of improvement: The power law applies by strategy within task. *Psychological Science, 9*, 1-8.

Ericsson, K. A., & Harris, M. S. (1990). Expert chess memory without chess knowledge. A training study. *Paper presented at the 31st Annual Meeting of the Psychonomic Society*, New Orleans, LA.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.

Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science, 8*, 305-336.

Freyhoff, H., Gruber, H., & Ziegler, A. (1992). Expertise and Hierarchical knowledge representation in chess. *Psychological Research, 54,* 32-37.

Fisk, A. W. & Lloyd, S. J. (1988). The role of stimulus to role consistency in learning rapid application of spatial rules. *Human Factors, 30,* 35-49.

Gobet, F. (1993). A computer model of chess memory. *Proceedings of 15th Annual Meeting of the Cognitive Science Society* (pp. 463-468). Hillsdale, NJ: Erlbaum.

Gobet, F. (1998). Memory for the meaningless: How chunks help. *Proceedings of the 20th Meeting of the Cognitive Science Society.* (pp. 398-403). Mahwah, NJ: Erlbaum.

Gobet, F., & Jansen, P. (1994). Towards a chess program based on a model of human memory. In H. J. van den Herik, I. S. Herschberg, & J. E. Uiterwijk (Eds.), *Advances in Computer Chess 7.* Maastricht: University of Limburg Press.

Gobet, F., & Simon, H. A. (1996a). Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin & Review, 3,* 159-163.

Gobet, F., & Simon, H. A. (1996b). Recall of random and distorted positions: Implications for the theory of expertise. *Memory & Cognition, 24,* 493-503.

Gobet, F., & Simon, H. A. (1996c). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31,* 1-40.

Gobet, F. & Simon, H. A. (1998). Expert chess memory: Revisiting the chunking hypothesis. *Memory, 6,* 225-255.

Gobet, F. & Simon, H. A. (in press). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science.*

Holding, D. H. (1985). *The psychology of chess skill.* Hillsdale, NJ: Erlbaum.

Hyötyniemi, H., & Saariluoma, P. (1998). Simulating chess players' recall: How many chunks and what kind can they be? *Proceedings of the Second European Conference on Cognitive Modelling* (p. 195-196). Nottingham: University Press.

Kintsch, W. (1998). *Comprehension. A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Laine, T., Hyötyniemi, H., & Saariluoma, P. (1998). Foundations of simulative theorizing. *Proceedings of the 13th European Conference on Artificial Intelligence* (p. 109-113). London: Wiley.

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and law of practice. In J. R. Anderson (Ed.), *Cognitive skill and their acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Richman, H. B., Staszewski, J. J., & Simon, H. A. (1995). Simulation of expert memory with EPAM IV. *Psychological Review, 102*, 305-330.

Saariluoma, P. (1984). *Coding problem spaces in chess: A psychological study.* Turku: Societas Scientiarum Fennica.

Saariluoma, P. (1989). Chess players' recall of auditorily presented chess positions. *European Journal of Cognitive Psychology, 1,* 309-320.

Saariluoma, P. (1995). *Chess players' thinking.* London: Routledge.

Saariluoma, P. & Laine, T. (in press). Novice construction of chess memory. *Scandinavian Journal of Psychology.*

Shiffrin, R. M. (1996). Laboratory experimentation on the genesis of expertise. In K. A. Ericsson (Ed.), *The Road to Excellence*. Mahwah, NJ: Erlbaum.

Simon, H. A., & Barenfeld, M. (1969). Information processing analysis of perceptual processes in problem solving. *Psychological Review*, *76*, 473-483.

Simon, H. A., & Gilmartin, K. J. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, 5, 29-46.

Zhang, G., & Simon, H. A. (1985). STM capacity for Chinese words and idioms: Chunking and acoustical loop hypothesis. *Memory and Cognition, 13,* 193-201.

Table 1

*Amount of variance accounted for by a logarithmic function (y = a + log [x]) and a power function (y = a * x^b) in Saariluoma and Laine's data and in CHREST simulations.*

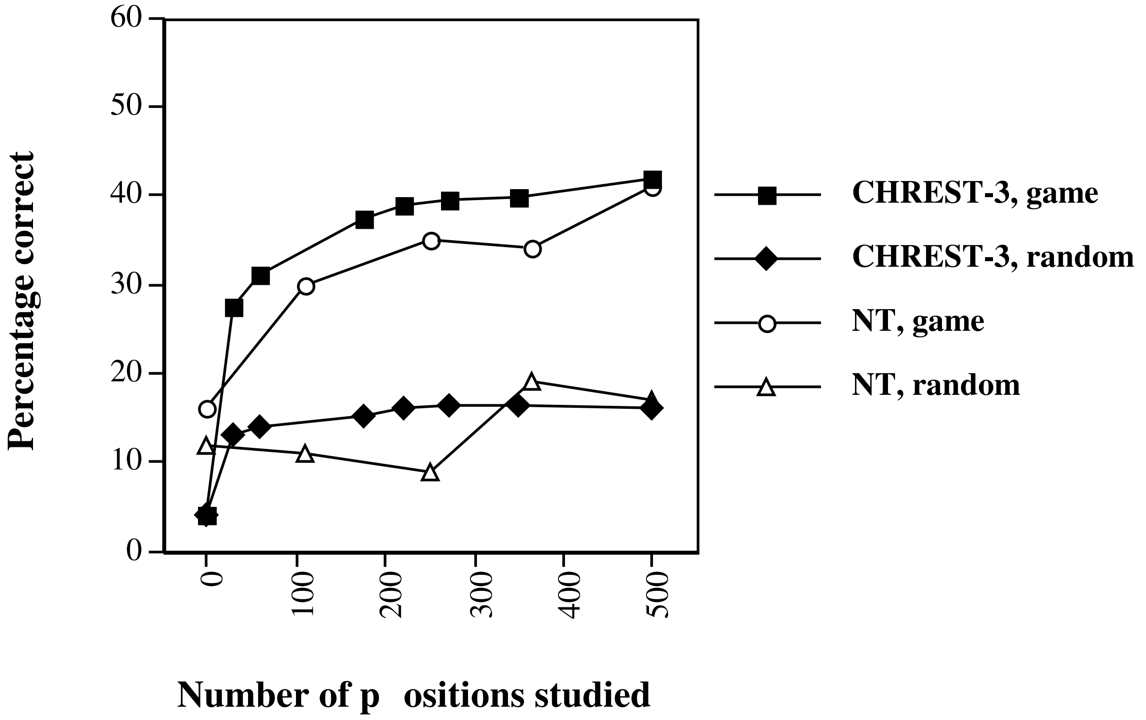
	Function	
Data source	Logarithmic	Power
NT, game	0.782	0.811
NT, random	0.475	0.438
CHREST-3, game	0.988	0.980
CHREST-3, random	0.936	0.935
MQ, game	0.948	0.949
MQ, random	0.477	0.417
CHREST-5, game	0.993	0.989
CHREST-5, random	0.922	0.925

Figure captions

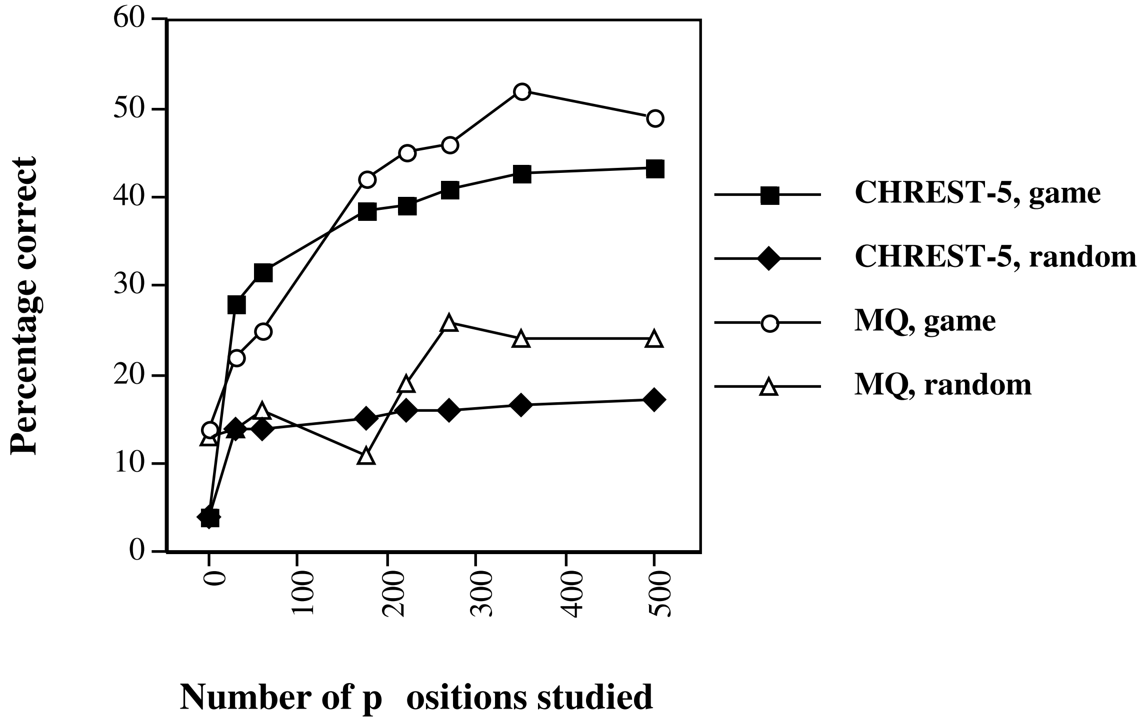
Figure 1. Percentage correct as a function of the number of positions studied. Upper panel: results for NT and for CHREST with a study time of 3 minutes per position. Lower panel: results for MQ and for CHREST with a study time of 5 minutes per position.

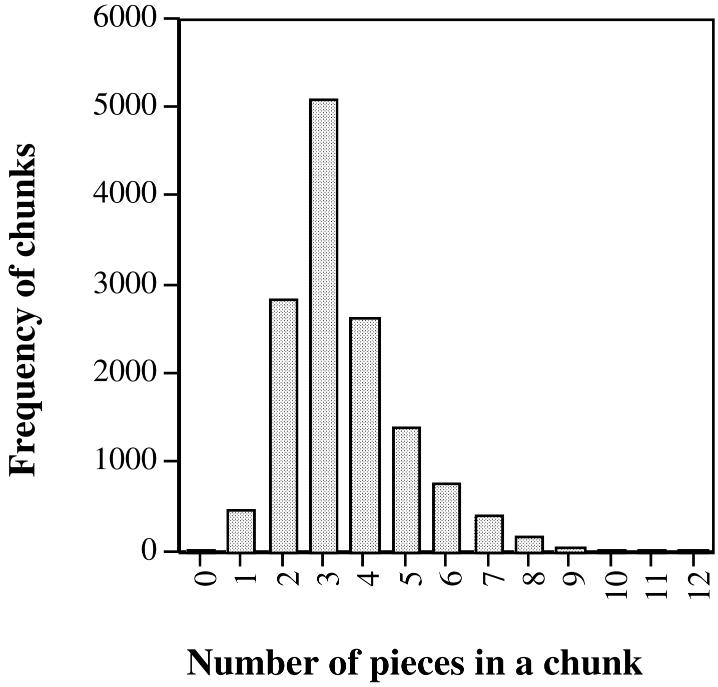
Figure 2. Distribution of chunks acquired by CHREST after 500 positions, with a study time of 5 minutes per position.

Results and simulations of NT



Results and simulations of MQ




6000
5000
4000
3000
2000
1000
0
Frequency of chunks
0
1
2
3
4
5
6
7
8
9
10
11
12
Number of pieces in a chunk